

The Developmental Sequence of Cognition
as Revealed by Wrong Answers

J. C. POWELL
University of Windsor

J. C. POWELL

University of Windsor

The Developmental Sequence of Cognition as Revealed by Wrong Answers

Gorham's Proverbs Test (1957) was administered to 550 children from three different schools in grades 3 to 8 inclusive. The highest and the lowest scoring 16% were interviewed to determine the reasons for the wrong answers given. Wrong answers were then grouped by a two-stage cluster analysis procedure and the resulting clusters used to develop subtest scores. The correlation matrix (of right and wrong answers) was reordered into a Simplex pattern. The resulting pattern clearly supported the possibility of a developmental sequence. The clusters were then interpreted, using the transcribed reported reasons for answers as a guide. These interpretations produced a logical description of the apparent developmental sequence when placed in the same order as revealed by the Simplex arrangement. A comparison was made between this order and the hierarchical sequence found elsewhere with mature adults (Powell, 1970). The rank order correlation between the orders of the common elements in the two studies was .86, suggesting a strong relationship and supporting the empirical validity of the order obtained. (Dr. Powell is Assistant Professor in the Faculty of Education at the University of Windsor.)

About fifteen years ago, the present author noticed a tendency for consistency in the selection of wrong answers on multiple choice tests. Specifically, a number of students tended to choose several wrong answers in common from several different items. Also, asking specific students why they chose these wrong answers often assisted in the planning of instruction.

These observations led to the question, "Are wrong answers really blind guesses as is commonly held?" A careful study using Gorham's Proverbs Test (1957) with adults revealed that this assumption is false (Powell, 1968).

Within the logic of test analysis, it is not sufficient that wrong answers be systematically selected. If answers are either right or wrong, then knowing that an answer is wrong adds no new information to the observation than an answer is not right. Such a "linear dependency" assumption is commonly held concerning wrong answers on multiple choice tests. This linear dependency question was explored (Powell & Isbister, 1974), using 387 adults and a different test, and was found to be false. With more than one alternative from which to choose, at least some wrong answers load on factors which contain no right answers; that is, some are not linearly dependent.

This latter discovery led to the question, "Do wrong answers contain achievement information?" The next study in the series (Powell, 1970) used the same test as in Powell and Isbister (1974). This study with 277 mature adults explored this question, with three conclusions:

1. Wrong answers may be better predictors of present and future performance than right answers (i.e., they contributed more of the variance, using stepwise multiple regression, to the prediction equation). Performance is defined in terms of obtained scores on course related achievement tests.
2. Wrong answers appear to be ordered in a hierarchy which parallels Bloom's *Taxonomy* (1956) and which influences the level of functioning of the questions in which they are embedded as distractors.
3. When the results of wrong and right answers are combined, the most stable (in cross-validation terms) prediction of present and future success is obtained.

Research up to this present study into the nature of behavior related to wrong answers may be summarized as follows:

1. The common assumption that a person writing a multiple choice test always either knows the answer or guesses blindly is false.
2. The related assumption that all wrong answers are linearly dependent upon the right answers is false.
3. The assumption that wrong answers contain no useful information (in a predictive sense) about achievement is false.

These conclusions imply that answer selection is much more complex than a know-guess dichotomy would seem to suggest. Do people choose answers based upon some systematic aspect of their personality? The "yes" answer to this question can be supported if a developmental sequence of wrong answer selection can be uncovered. It is to this latter issue, the exploration for a developmental sequence, that this present study turns.

Method

Instrument

Gorham's Proverbs Test (1957), in its multiple choice version, involves 40 proverbs for each of which four possible translations are offered. Although it also has clinical uses, its primary value with normal subjects is as a short, broad-spectrum, reading comprehension test. Norms are offered in the manual for the range of grade 5 through adult for the Abstract scale. Norms are not given for the Concrete scale. The wrong answer interpretations reported here were developed in the present study.

Procedure

This test was administered to 550 children in grades 3 to 8 inclusive (approximately 90 children in each grade) from three different schools. The five highest scoring and the five lowest scoring children from each grade level of each school were interviewed to determine the reasons they had for the selection of the wrong answer. This interview followed the administration of the tests which were scored immediately so that the children could be brought to the examining room usually within minutes of the completion of the test. Two trained interviewers were employed in each school. In no case were the reasons collected more than two hours later than the test administration. This procedure was used to minimize the effects of forgetting.

A frequency count of the selection of each of the four alternatives partitioned by age was produced. The proportion of selection of each wrong alternative by age was calculated. It was observed that particular alternatives tended to have the highest proportion of selection at particular age levels. These were grouped together and further analyzed so that alternatives which were significantly correlated were placed into definitive clusters. A unique group was also generated from some alternatives which were bimodal in their highest proportions.

Thus age-level related groups of wrong alternatives were placed into homogeneous clusters. Those clusters with five or more members were treated as subtests and the answers of all 550 children were rescored. Those which formed clusters of less than five were dropped from the analysis. Thus the present study accounts for 82 of the 101 possible wrong answers. The resulting 14 profile scores (including the two right answer scales) were correlated and the resulting correlation matrix was rearranged into the Simplex order as defined by Guttman (1944).

These clusters were also classified using the reported reasons for selection, so that if a hierarchy emerged, its developmental properties could be examined.

Results

The first question considered was whether or not a hierarchical ordering was evident. Table 1 gives the results of this analysis. The categories are coded by age (10y stands for 10 years old). If more than one subscore set was present for a particular age, these were assigned an arbitrary number (8y4 stands for the fourth subset among the 8-year-olds) to distinguish among them.

It is evident from the arrangement of subtest scores in Table 1 that a clearly definable age-dependent hierarchy was found among these data. In fact, not a single age order displacement occurs. The majority of correlations are significant ($p < .05$ for $r = .083$) and these clearly support the hierarchy. Most of the significant correlations are low, once again providing evidence non-supportive of the "linear dependency" assumption.

The reasons given within each cluster for the selection of the particular alternatives in it tended to be logically equivalent between 50 and 60% of the reasons reported.

Proceeding along the hierarchy as given by the Simplex order from low to high, the interpretation given in Table 2 emerged.

TABLE 1
SIMPLEX^a ORDER OF THE CORRELATIONAL ANALYSIS
OF ALL VARIABLES

Right Answer Scales	Wrong Answer Scales	14y+	13y	12y	Bimodal	11y	10y	9y2	9y1	8y3	8y1	Concrete	8y2	8y4
ABSTRACT		<u>.25</u> ^c	.06	.00	<u>-.20</u>	.05	<u>-.22</u>	<u>-.23</u>	<u>-.44</u>	<u>-.46</u>	<u>-.48</u>	<u>-.73</u>	<u>-.60</u>	<u>-.31</u>
	14y+ ^b		.07	.01	<u>-.01</u>	<u>-.09</u>	<u>-.16</u>	<u>-.11</u>	<u>-.18</u>	<u>-.14</u>	<u>-.16</u>	<u>-.22</u>	<u>-.15</u>	<u>-.16</u>
	13y			<u>-.07</u>	.04	.05	<u>-.01</u>	<u>-.08</u>	<u>-.10</u>	<u>-.10</u>	<u>-.16</u>	<u>-.18</u>	<u>-.14</u>	<u>-.05</u>
	12y				.05	.06	.01	<u>-.16</u>	<u>-.08</u>	<u>-.05</u>	.00	<u>-.02</u>	<u>-.06</u>	<u>-.04</u>
	Bimodal					.01	.09	.02	<u>-.00</u>	<u>-.01</u>	.02	<u>-.04</u>	.04	<u>-.05</u>
	11y						<u>-.14</u>	.06	<u>-.08</u>	<u>-.10</u>	<u>-.24</u>	<u>-.27</u>	<u>-.24</u>	<u>-.07</u>
	10y							<u>.17</u>	.03	<u>-.01</u>	<u>-.09</u>	<u>-.10</u>	<u>-.01</u>	<u>-.08</u>
	9y2								.05	<u>-.06</u>	.06	.06	<u>.12</u>	<u>-.02</u>
	9y1									<u>.15</u>	.02	<u>.16</u>	<u>.23</u>	.04
	8y3										<u>.36</u>	<u>-.14</u>	<u>.26</u>	.06
	8y1											<u>.37</u>	<u>.26</u>	<u>.22</u>
CONCRETE													<u>.39</u>	<u>.21</u>
	8y2													<u>.16</u>

^a The simplex arrangement involves ordering correlations in increasing magnitude vertically upward and horizontally to the left.

^b To give a reasonable group size, children of 14, 15, and 16 years of age were combined into one group.

^c Underlined correlations are significant at $p \leq .05$.

TABLE 2
INTERPRETATION OF WRONG ANSWER CLUSTERS

General Stages	Wrong Answer Scales	Interpretation	No. of Alternatives in Scale
Personalized interpretations	8y4	Partial translations	5
	8y2	Isolated responses	15
	8y1	Redefining terms	9
	8y3	Personalized answers	7
Literalized interpretations	9y1	Word associations	8
	9y2	Literal reductions	5
	10y	Literalizations	9
Figurative interpretations	11y	Oversimplifications	6
	Bimodal	Irrelevancies	10
	12y	Transpositions	4
	13y	Overgeneralization	6
	14y+	Simplifications	7

The names for these wrong answer categories were assigned as a result of the interpretive analysis conducted in the study.

These 12 classes of response use 82 of the 101 possible wrong responses. There are 19 answers classified as concrete correct answers and 40 as

abstract correct answers. Thus this procedure uses 141 of the 160 (4×40) alternatives on the test, or 88% of all possible answers.

Three general classes of wrong answer emerged. First, personalized interpretations all involved the typical egocentric behavior described by Piaget in which the child responds within the framework of her or his own limited personal experience and is unable to entertain externalized reference points. Second, literalized interpretations all involve external reference to the denotive rather than the connotive meanings of the words used. Third, the child in this stage shows greater or lesser facility in the use of the connotative or figurative meanings of words.

A detailed description of the basis for the interpretation of the types of error identified is shown in Table 3. These are given in the same order as in Table 2. Each description contains (1) the name assigned to this type of response by the present researcher; (2) a behaviorally based definition for this type of response derived from the reasons reported by the examinees within each type; (3) an illustration of the behavior with a typical reason

TABLE 3(a)
TYPES OF ERROR IDENTIFIED IN STUDY
PART I: PERSONALIZED INTERPRETATIONS

Wrong Answer Scale	8y4	8y2	8y1	8y3
<u>Classification</u>				
Interpretation Assigned	Partial translation	Isolated responses	Redefining terms	Personalization
Behavioral Definition	Only part of the proverb is translated into personal experience.	Explanation of answer selection related to distractor only.	The meaning of the proverb is changed by using alternate meanings for key words.	The child relates both the stem and the foil to the same personal experience, often correctly.
<u>Illustration</u>				
Proverb	THE HOT COAL BURNS, THE COLD ONE BLACKENS	THERE'S MANY A SLIP TWIXT (between) THE CUP AND THE LIP	GOLD GOES IN AT ANY GATE EXCEPT HEAVEN'S	DON'T THROW GOOD MONEY AFTER BAD
Translation Chosen	Leave dangerous things alone.	Don't talk too much while eating.	Anyone would take money.	Don't waste money.
Reason Given	You might get burned badly.	You'll bother people if you talk too much.	I think people should work for their MONEY. (i.e., take = steal).	Because if you waste all your money you will end up poor.
<u>Overall Consistency (Per cent)</u>				
	54.9	56.7	59.4	50.9

TABLE 3(b)
TYPES OF ERROR IDENTIFIED IN STUDY
PART II: LITERALIZED INTERPRETATIONS

Wrong Answer Scale	9y1	9y2	10y
<u>Classification</u>			
Interpretation Assigned	Word associations	Literal reduction	Literalization
Behavioral Definition	The child makes a correct association between a word or two in the stem and in the foil and ignores the rest of the problem.	The child extracts most but not all of the literal meaning of the basic communication.	The child translates both stem and foils into a literal statement in his/her own words and matches these literal statements for his answer.
<u>Illustration</u>			
Proverb	THE MORE COST, THE MORE HONOR	DON'T SWAP (TRADE) HORSES WHEN CROSSING A STREAM	ALL IS NOT GOLD THAT GLITTERS
Translation Chosen	Good things have to be paid for in some way.	Don't try something until you're able to do it.	Some things may fool you.
Reason Given	Good things cost more.	You might fall in.	A lot of stones look alike.
<u>Overall Consistency</u> (Per Cent)			
	53.7	57.8	57.1

reported; and (4) the proportion of all respondents in this type whose reported reasons are consistent with the behavioral description. Table 3 emphasizes the three stages in development uncovered among these data in this study.

Problems of order occur at both ends of this sequence. Since the study was conducted in May, the 8-year-old group represents the youngest grade 3 students, about half of whom had become 9 years old during the year. With no reference point lower, it is difficult to establish the order in which the four 8-year-old subscores should be reported. Similarly, since the 14-, 15-, and 16-year-olds are the oldest grade 8 students, the manner in which Oversimplification (with the 11-year-olds) and Simplification (with 14-year-olds) is different is not entirely clear.

It is possible, however, to produce a general two-way classification for these wrong answers. This system is presented in Table 4 and presents the suggestion that particular error types may recur within the general framework of the modifications in orientation that the different stages seem to imply.

When the classified sequence found in this study is rank order correlated with the equivalent categories in the hierarchy found elsewhere among mature adults (Powell, 1970), the correlation is .86. It is evident from this

Developmental Sequence of Cognition

TABLE 3(c)
TYPES OF ERROR IDENTIFIED IN STUDY
PART III: FIGURATIVE INTERPRETATIONS

Wrong Answer Scale	11y	Bimodal	12y	13y	14y+
<u>Classification</u>					
Interpretation Assigned	Over-simplification	Irrelevancy	Transposition	Over-generalization	Simplification
Behavioral Definition	The child correctly translates most, but not all, of the figurative meaning.	The child chooses a true statement which is unrelated to the generally accepted meaning of the passage.	The child changes substantially the figurative meaning of the proverbs.	The child chooses an answer which goes beyond the scope of the usual figurative meaning.	The child correctly interprets most but not all of the usual figurative meaning of the passage. To distinguish this class of responses from the 11y one, it is necessary to think in terms of more abstract and/or nearer correctness for this class.
<u>Illustration</u>					
Proverb	RICHES SERVE A WISE MAN BUT COMMAND A FOOL	IT NEVER RAINS BUT IT POURS	THE GRASS IS ALWAYS GREENER IN THE OTHER FELLOW'S YARD	TOO MANY COOKS SPOIL THE BROTH	SPEECH IS THE PICTURE OF THE MIND
Translation Chosen	Don't let money go to your head.	A little is as bad as a lot.	Don't stay in one place to do something.	You'll have too many ideas.	Words paint pictures in your mind.
Reason Given	Money can't (shouldn't) run your life.	If you do something wrong, it's still wrong, even if you do just a little thing or something real bad.	Because if you stay in one place all the time you'll never know what the other place is like.	If you have too many people, they won't be able to decide what to do.	When you're reading, sometimes you can picture what the author's telling you.
<u>Overall Consistency</u> (Per Cent)					
	57.1	59.4	52.9	53.7	62.7

TABLE 4
DEVELOPMENTAL SEQUENCE OF WRONG ANSWERS

General Strategy		Level of Abstraction of Interpretation							
		Personalized			Literalized			Figurative	
Fragmentation		Isolated responses			Word associations			Oversimplification	
Reduction		Partial translation			Literal reductions			Simplification	
Rearrangement		Redefined terms			Transpositions			(Inversions)	
Extension		(Continuations)			(Extrapolations)			Overgeneralizations	
Transition		[Concrete answers]			Irrelevancies				
Balancing		Personalities			Literalizations			[Abstract answers]	
Age Scale		7	8	9	10	11	12	13	14

() Not found in this study.

[] Correct answer subtests.

observation that there is a high level of consistency in the findings between these two independent studies which used different instruments and different age groups. There was also observed to be a perfect agreement for equivalent classes with the error sequence reported in the manual, *Raven's Coloured Progressive Matrices* (1956). Thus this sequence would seem largely to recapitulate the findings of other independent studies on different populations.

Conclusions

The major characteristics of these newly observed classes of behaviors are fairly evident from the findings reported in this study. Children seem to move from a personalizing stage where they relate all problems to their personal experience through a stage when they take meanings literally before they begin to deal figuratively with meaning in any systematic manner.

It is therefore reasonable to report that this study has discovered a previously unreported aspect of the developmental sequence of behaviors in that wrong answers seem to parallel the observations made by Piaget and Vygotsky, as reported by Stewin and Martin (1974) among others, with respect to the general learning sequence.

Since this study was cross-sectional rather than longitudinal, there is no way of telling if all children go through all these phases. Hence the apparent recurring cycle reported in Tables 3 and 4 may not actually exist. If these recurrences do exist, it may imply that there may be specific thinking styles which particular children characteristically use. The possibility would explain the predictive power of wrong answers observed elsewhere (Powell, 1970). Also, there may be particular transitions through which a child passes and optimal progress takes place.

In any case, the discovery that a type of behavior not usually considered in research studies (namely wrong answers) follows a similar developmental sequence to the one commonly observed in other phases and stage-oriented research has interesting implications for further research. The principal observation here suggests that the way a person interprets and attacks a problem determines the answer selected. In this case, learning may not be a simple linear accumulation of information and skills as suggested by Gagné (1965).

References

- Bloom, B. S. *Taxonomy of educational objectives: Handbook I: Cognitive domain*. New York: David MacKay, 1956.
- Gagné, R. M. *The conditions of learning*. New York: Holt, Rinehart and Winston, 1965.
- Gorham, D. R. *The proverbs test*. Missoula, Montana: Psychological Test Specialists, 1957.
- Guttman, L. A basis for scaling qualitative data. *American Sociological Review*, 1944, 9, 139-150.
- Powell, J. C. The interpretation of wrong answers from a multiple choice test. *Educational and Psychological Measurement*, 1968, 28(2), 403-412.
- Powell, J. C. *Achievement information from wrong answers*. (Short title). Unpublished doctoral dissertation, The University of Alberta, Edmonton, Alberta, May, 1970.

Developmental Sequence of Cognition

- Powell, J. C., & Isbister, A. B. A comparison between right and wrong answers on a multiple choice test. *Educational and Psychological Measurement*, 1974, 34(3), 499-509.
- Raven, J. C. *Guide to using the Coloured Progressive Matrices*. London: H. W. Lewis, 1956. (Reference Printing of 1973 pages 21 and following.)
- Stewin, L. L., and Martin, J. The developmental stages of L. S. Vygotsky and J. Piaget: A comparison. *Alberta Journal of Educational Research*, 1974, 20(4), 348-362.

